Before the
Federal Communications Commission
Washington, D.C. 20554

In the Matter of	)	
Closed Captioning of Video Programming	)	CG Docket No. 05-231 ET Docket No. 99-254

COMMENTS OF THE
WGBH NATIONAL CENTER FOR ACCESSIBLE MEDIA
(NCAM)
November 24, 2010

Submitted By:
Larry Goldberg and Marcia Brooks
WGBH National Center for Accessible Media
One Guest Street
Boston, MA 02135

The WGBH Educational Foundation's National Center for Accessible Media (NCAM) hereby submits comments on the Commission's Pleading Cycle to refresh the record in the proceeding noted above concerning the Commission's Closed Captioning Rules.

Summary:

1. The FCC has asked for comment on whether the Commission should establish quality standards for non-technical aspects of closed captioning, including the accuracy of transcription, spelling, grammar, punctuation and caption placement, what the adoption of such standards would cost to programmers and distributors, whether the captioning pool consists of an adequate number of competent captioners to meet a non-technical quality standard mandate, and whether different captioning quality standards should apply to live and pre-recorded programming.

2. The FCC has asked for comment to refresh the record regarding the need for mechanisms and procedures, over and above the "pass through" rule, to prevent technical problems from occurring and to expeditiously remedy any technical problems that do arise, including current and proposed obligations for video programming distributors to monitor and maintain their equipment and signal transmissions.

3. The FCC has asked for additional comment on whether to establish specific per violation forfeiture amounts for non-compliance with the captioning rules, and if so, what those amounts should be, and whether video programming distributors (VPDs) should be required to file closed captioning compliance reports.

4. Since filing comments on this proceeding on November 10, 2005, the WGBH National Center for Accessible Media (NCAM) has conducted significant research and development that now advances the Commission's ability to establish quality standards. NCAM believes the Commission should indeed establish standards for non-technical quality of closed captioning.

Introduction

5. The WGBH Educational Foundation is one of the country's leading public broadcasters and has long considered one of its central missions to be increasing access to media for people with disabilities.

6. WGBH's commitment to accessible information began in 1971 through establishment of The Caption Center, the world's first captioning agency, to produce captions for TV programs so that deaf and hard-of-hearing viewers could gain equal access to those programs. Today, The Caption Center is part of WGBH's Media Access Group and produces captions and subtitles for every facet of the television and home video industry. The Media Access Group additionally services the theatrical film industry, museums and theme park attractions.

7. The WGBH Media Access Group also houses WGBH's Descriptive Video Service ® (DVS ®) which makes television programs and movies accessible to people who are blind and visually impaired. WGBH developed DVS in 1990 and continues to lead the world in creating accessible electronic media for people with disabilities.

8. The WGBH National Center for Accessible Media was founded in 1993 to build on WGBH's knowledge base in the field of access technologies. NCAM is a research and development facility

dedicated to addressing barriers to media and emerging technologies for people with disabilities in their homes, schools, workplaces, and communities.

9. These comments expand upon comments The WGBH National Center for Accessible Media previously submitted in November 2005 on the Commission's Notice of Proposed Rule Making concerning the closed captioning of television programs.

Non-technical Quality Standards for Closed Captioning – The Marketplace Has Still Not Corrected Problems

10. Caption errors continue to be pervasive, especially as the use of Automatic Speech Recognition (ASR) – a technology not ready to be used for real-time captioning – is becoming more common. The lack of a common way to measure accuracy may have held back establishment of quality requirements in the past, but with newly developed technology created by WGBH/NCAM's innovators with significant input from caption users, deaf education experts, and with measurement parameters developed by the

National Institute of Standards and Technology (NIST)[1] and National Court Reporters Association (NCRA)[2], the FCC can now set fair levels of expected performance.

11. NCAM is developing a prototype automated caption accuracy assessment system that will identify, rank and report on the frequency and severity of caption errors through its Caption Accuracy Metrics project (funded by the National Institute on Disability and Rehabilitation Research, U.S. Department of Education, #H133G080093-10)[3].

Current State of Caption Accuracy Measurement

12. Accuracy measurements are traditionally based on the model used at the National Institute of Standards and Technology (NIST). This approach identifies the differences between a test transcript (in this case, a caption text file) and a clean reference transcript, often called the "ground truth" transcript, which accurately reflects

[1] NIST: http://www.nist.gov
[2] NCRA: http://www.ncraonline.org/
[3] Caption Accuracy Metrics project: http://ncam.wgbh.org/invent_build/analog/caption-accuracy-metrics

what was spoken. The two transcripts are aligned and errors are categorized as:

- Substitutions – words in the test transcript that are different from the reference transcript;
- Deletions – words that are in the reference transcript but are omitted from the test transcript; and
- Insertions – words that are added to the test transcript but are not in the reference transcript.

The total number of these errors is divided by the total word count of the reference transcript to calculate a Word Error Rate. An accuracy rate is 100% minus the error rate. Accuracy rates for most caption text range from 85 to 95% by this measure, with lower accuracy usually due to more extensive deletion of text.

Caption agencies have used a different approach to error reporting for live stenocaptioning. Court reporting software used by most captioners identifies “untranslates” – words that do not have a match in the stenocaptioner’s dictionary. These reflect a portion of the substitutions that would be found in the caption file but they do not

typically identify deletions or insertions. Accuracy rates for caption text by this measure usually fall in the 97-99% range.

Quality Standards Informed by NCAM Research and Development

13. Technical development to date for the Caption Accuracy Metrics project demonstrates a proof of concept that text-based data mining and automatic speech recognition technologies can produce meaningful data about stenocaption accuracy that meets the need for caption performance metrics.

14. Further, it is now possible to quantify the severity of specific caption error types and to specify the degree to which each error type makes a caption hard to follow, using data from a national consumer research web-based survey the Caption Accuracy Metrics project conducted in Spring 2010 that yielded over 350 responses from caption viewers. Caption viewers were presented with actual caption error samples representing 17 different error types, and they ranked the severity of each error type. The survey results provide valuable data about how to rank the severity of the

17 types of errors evaluated through this survey. The summary consumer research report will be available in December 2010 at the Caption Accuracy Metrics project website.

15. Combining the research and development as noted above, it is now possible to generate an accuracy report per program that estimates the level of caption accuracy using Automatic Speech Recognition (ASR). This process occurs after the real-time captioned program is broadcast and is not utilizing ASR to generate captions.

Further Definitions of Caption Accuracy

16. NCAM developed a caption error ontology that identifies 17 caption error types sub-categorized by the major three error types identified by the National Institute of Standards and Technology (insertions, substitutions and deletions), and assigns a severity ranking informed by the consumer research data. This ontology addresses many of the questions identified by the FCC such as spelling, grammar, and punctuation. The ontology and the severity ranking for each error type are expanded upon in the Caption

Accuracy Metrics survey report, which notes there is a wide range of error types in real time captioning and they are not all equal in their impact to caption viewers. Treating all substitution and deletion errors the same does not provide a true picture of caption accuracy. The least offensive errors were judged to be simple "substitutions" like the wrong tense and punctuation; however, substituting pronouns and/or nominals for proper names were also judged to significantly impact viewers' understanding.

17. In September 2010, The Caption Accuracy Metrics project convened a technical review panel consisting of many of the major stakeholders in caption quality (including broadcast and cable television networks, caption vendors, deaf education experts, and the National Court Reporters Association). There was wide consensus that each sector would fully support defined caption quality standards, but only if there is full and equitable compliance across the range of industry stakeholders. NCAM believes it is the FCC's role to define and ensure compliance with caption quality standards.

18. NCAM believes that the Commission should include and define caption placement requirements in its caption accuracy standards. Through research and development NCAM conducted for its Access to Locally-Televised Onscreen information project (funded by the U.S. Department of Education, National Institute on Disability Research and Rehabilitation, grant #H133G070278)[4], NCAM developed a prototype system that demonstrates the ability to automatically resolve display conflicts between captions and on-screen graphics. By developing methods of prioritizing text and graphics messages within automated display systems, the system automatically relocates closed captions so they are not obscured by emergency information (also known as "crawls") located on the screen. Note that the system also automatically translates the text in the emergency crawls to speech, for viewers who are blind or low-vision.

[4] Access to Locally Televised On-Screen Information http://ncam.wgbh.org/invent_build/analog/onscreen

Establishment of Reporting Requirements and Non-Compliance Forfeiture Amounts

19. The FCC has asked for additional comment on whether to establish specific per violation forfeiture amounts for non-compliance with the captioning rules, and if so, what those amounts should be, and whether video programming distributors (VPDs) should be required to file closed captioning compliance reports. NCAM believes that the Commission should establish and enforce VPD reporting requirements that are developed in parity as appropriate with other existing FCC reporting requirements where a structure to manage reporting requirements exists or has been defined (e.g., telecommunications industry network outage reports, etc.). Because the marketplace has not significantly corrected caption quality problems, and because the means by which to define and measure caption quality standards are being established, further examination of forfeiture amounts - perhaps tied to compliance reporting requirements - is recommended.

Cost of Adoption of New Caption Standards

20. From the Caption Accuracy Metrics technical review panel, which represents a wide range of stakeholders in caption quality, it is apparent that many video program distributors (VPDs) and captioning agencies are already monitoring caption quality to some degree, and in some cases service level agreements exist between television networks and their caption vendors. However, there is not a standard way to define or measure caption quality. Many panel members agreed that an automated system of caption quality monitoring would in many cases ultimately decrease the cost of monitoring caption accuracy and levels of service they are currently tracking through labor-intensive, manual means. If the Commission indeed sets caption quality standards, all stakeholders -- VPDs and caption agencies who are already tracking accuracy levels as well as those who do not currently have an established means to do -- so will be at an advantage, given the likelihood of having access to an automatic system to measure caption accuracy. The upfront costs of such a system are yet to be determined, but are likely to ultimately be a far more cost-effective option than manual monitoring and/or payment of

potential fines. VPDs further stand to benefit from an automatic system that can identify caption errors such as garbling caused by technical errors, which can help inform troubleshooting of the transmission equipment chain. Establishment of caption quality standards will also likely ease the significant burden on consumers to report caption quality issues, and therefore, also ease the burden on local television stations, the FCC and national consumer advocacy organizations in responding to complaints from viewers who rely on closed captioning for equal access to information.